李航统计学习：

第五章讨论部分：

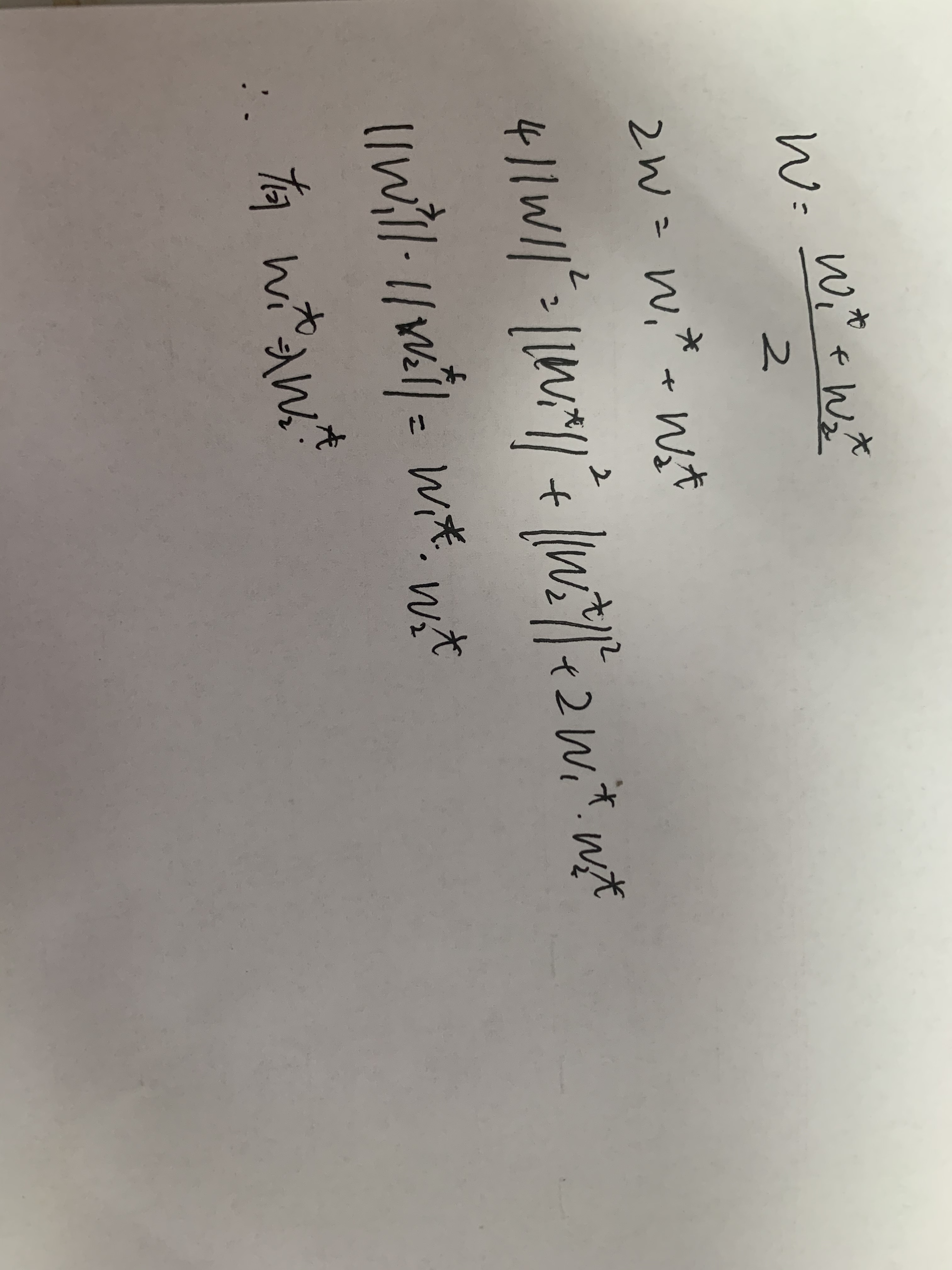
读书报告内容：

1. （必填）自己提出的问题的理解（罗列全部）：
2. 提出的问题1：

p117页，为什么从而有w\*1=lamda w\*2

讨论后的理解：

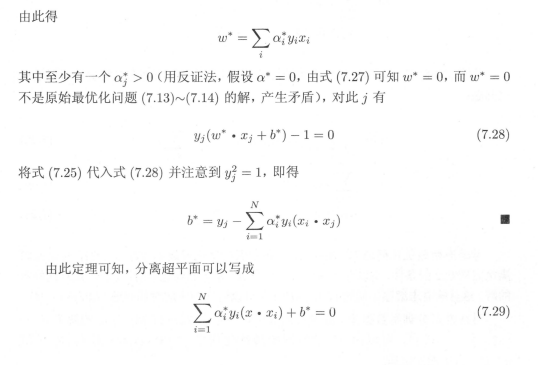
手写推导过程如下：



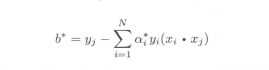
1. 提出的问题2：

p122页，分离超平面为什么可以写成这个式子？

讨论后的理解：



由方框式得：



又由上述得：



所以可以写成wTx+b=0的形式，b只是代数。

二、（必填）别人提出的问题的理解（选择几个问题罗列，并给出理解）：

问题3：p132，图7.6中0-1损失函数在间隔大于0的部分是什么样的。

自己的理解：

间隔大于0的部分，正例为0，负例为1.

问题4：p131，（7.58）为什么式中要有一个正则化项？

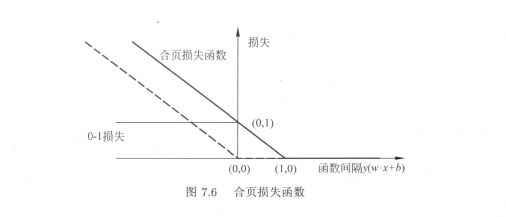
自己的理解：

因为正则化控制了||w||的大小，这样可以保证不同w的L2范式不会影响损失函数的计算。

问题5：P132页，为什么合页损失函数是0-1损失函数的上界？

自己的理解：

因为合页函数值始终比0-1损失函数的大，并且可导，经常用来代替0-1损失函数。



1. （必填）读书计划
2. 本周完成的内容章节：

（1）7.1到7.2完成并且梳理知识点，寻找问题，自己思考并且于小组会之前完成了自己的思考。

2、下周计划：第七章剩余.

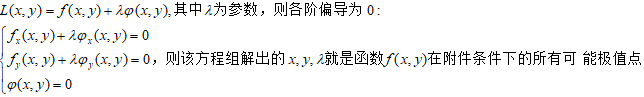
四、（选做）读书摘要及理解或伪代码的具体实现（读书摘要、伪代码的具体实现代码等可以写到这个部分）

拉格朗日乘数法原理证明：

参考知乎文章：马同学高等数学《如何理解拉格朗日乘子法》，下述为自己理解后复盘。

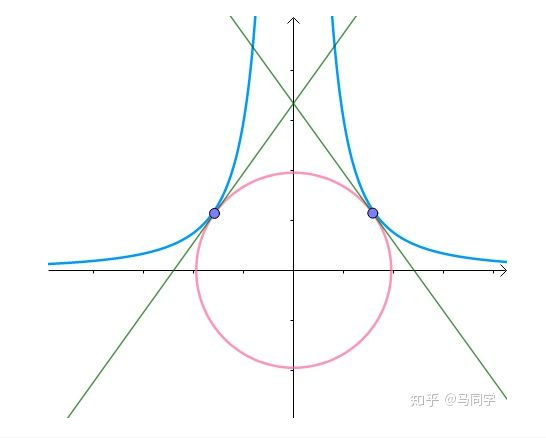
高数中学习的拉格朗日乘数法可以通过引入新的未知标量（拉格朗日乘数 [公式] ），直接求多元函数条件极值，不必先把问题转化为无条件极值的问题。

求函数f(x,y)在附加条件 [公式] 下可能的极值点，可以先做拉格朗日函数：



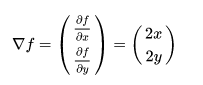
那么首先我们从二维进行分析

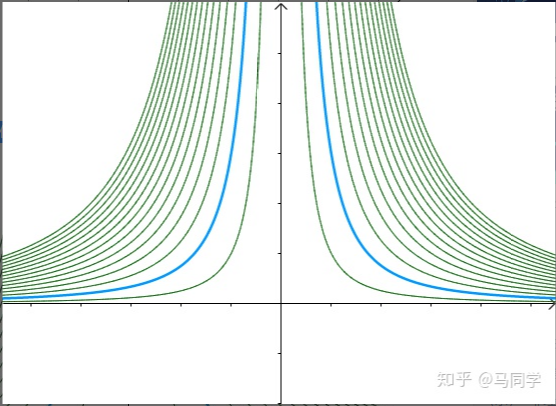
如图所示：求在约束条件x2\*y=3的情况下，求x2+y2的最值问题



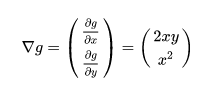
显然有：在极值点，圆与曲线相切。

下面我们引入等高线的概念，根据梯度的性质，梯度向量：

是等高线的法线，而另一个函数，g（x，y）=x2\*y 的等高线为下图所示，之前的曲线x2\*y=3就是其中值为3的等高线：



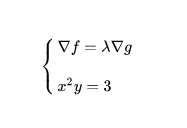
因此，必有这俩梯度向量相等，则为：



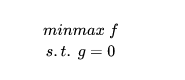
又因为有确切的证明可推导，梯度与等高线的切线垂直（可数学推导，运用求偏导的方法，此处忽略，高等数学知识）。

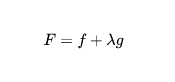
所以，其必有梯度向量平行，用数学符号表示为：1595823569(1)

故有联立方程：

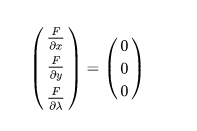


现在我们回头看拉格朗日乘子法的定义：



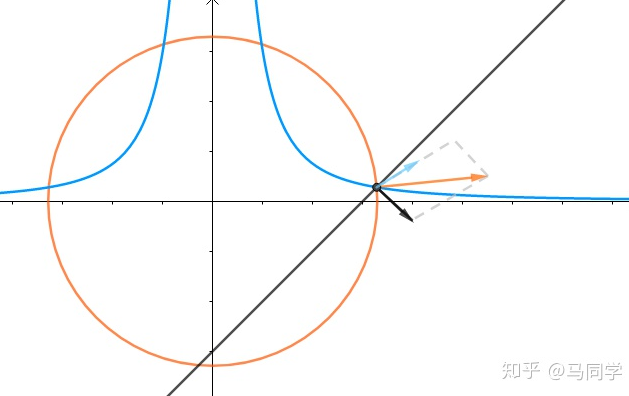
那么我们定义：

在对F的x，y，lamda求偏导，则有：

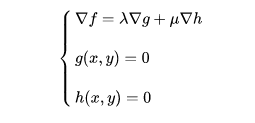


此处豁然开朗，突然理解了为什么之前学习高数，要构造拉格朗日F，再进行求偏导，之前学习只知其然不知其所以然，此时终于得其原理，感觉，数学真是妙不可言。

下面我们对拉格朗日乘子进行扩充增加多个约束条件，那么图为：



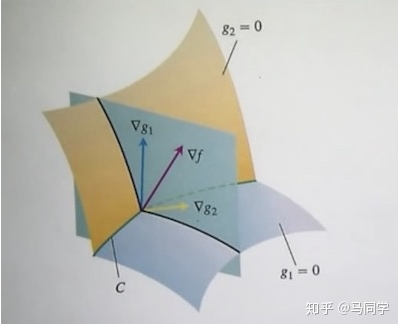
可得联立方程：



即可求解

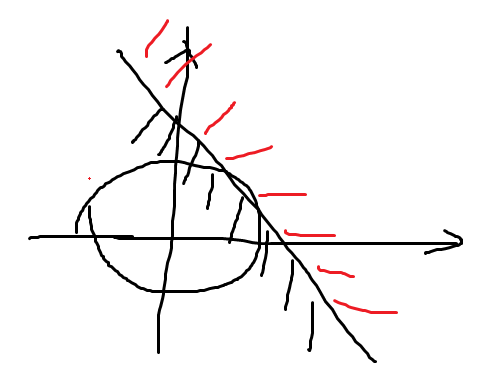
如果从高维向量高维超平面的角度讲直观上看

迪尔塔f必然在迪尔塔g1和迪尔塔g2形成的张开空间中，如图所示：



但是马同学原文这里就结束了，实际上，这只是在等式的情况下，而不等式的情况下，必然要考虑到lamda的梯度和不等式的梯度反向的情况下，其必有hj《=0，lamda》=0

这显然和我们所认知的支持向量在边界上所符合，在二维空间可表示为：

很显然，在向量同向的时候，即为图中的黑阴影部分，边缘边界上的点并不能作为支持向量，因为最小值点必定在原点（0，0），而反向时，即为红色阴影部分，才有作为支持向量。则有必要不充分条件lamda\*hj=0，这就是KTT条件里面的第三个约束条件。

综上，即为所证。